DOCUMENT RESUME

ED 191 887 TM 800 511

AUTHOR Fowler, William J., Jr.
TITLE Effects of School Characteristics upon Achievement Test Scores in New York State.
PUB DATE Apr 80
NOTE 31p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).
AVAILABLE FROM Bucknell University, Department of Education, Lewisburg, PA 17837 ($1.50).

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Achievement Tests; Elementary School Mathematics; Elementary Secondary Education; Grade 3; Grade 6; Grade 9; *Institutional Characteristics; Predictor Variables; Reading Achievement; *School District Spending; *Scores; Secondary School Mathematics; Socioeconomic Status
IDENTIFIERS *New York; *Pupil Evaluation Program

ABSTRACT

The effects of school characteristics upon achievement test scores in New York State were studied. Data, composed of the 1975-76 Consolidated Data base and Finance Tapes for all 705 school districts in the state, were supplied by the New York State Department of Education. Among the 24 variables of interest were: state pupil evaluation tests of reading and mathematics for all districts in grades 3, 6, and 9; and Regent's tests and school characteristics, such as per pupil expenditure and average teacher degree level. The data file was merged with the 1970 census data for New York school districts; occupational status scores, income, and education scores for 642 matched-districts were also used. The computer program, Statistical Package for the Social Sciences, was used to obtain means, frequencies, and skewness of the variables of interest. Results generated two conclusions: (1) a puzzling, inconsistent pattern between achievement test scores and school characteristics exists; and (2) additional expenditures per child and other quality-related measures raised the test scores of the poorest performing students. This study does not consider the range of standard deviations associated with school district attributes. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED191887

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EFFECTS OF SCHOOL CHARACTERISTICS UPON ACHIEVEMENT TEST SCORES IN NEW YORK STATE

by

William J. Fowler, Jr.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

W. Fowler
W. Iatt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A paper presented to the American Educational Research Association Annual Meeting in Boston, Massachusetts

April 8, 1980

TM 800511

TABLE OF CONTENTS

		Page
1.	Introduction	1
2.	Previous Findings	2
3.	Methodology	8
4.	Results	15
5.	Discussion	18

1. Introduction

The quintessential argument of those involved in the school finance reform movement is that existing disparities in per pupil expenditures should be reduced, and the ability to raise revenues equalized, because those differences reflect degrees of quality, rather than regional price differentials. Prior to the Coleman Report (1966), educators assumed that higher expenditures per pupil resulted in the acquisition of higher quality school district staffs and facilities, which produced superior achievement test scores. Coleman, however, demonstrated that school inputs did not necessarily affect school outputs, particularly achievement test scores, generating much debate. Extensive reanalysis of his data, adjusted for presumed methodological flaws, confirmed his initial findings (Armor, 1972). Questions concerning the accuracy of the data were allayed by Jencks (1972).

Other studies of school effectiveness [Averch, (1972), Suppes and Wells (1974), Gutherie (1970), Bargen and Walberg (1977)] have demonstrated contradictory findings. Typically, researchers in such studies sought to find a relationship between student achievement and per pupil expenditures or teacher characteristics such as experience, age, degree level, recency of education, salary, or building characteristics such as school size and numbers of pupils. The best of these studies controlled for socioeconomic status and a few removed the influence of previous achievement--and fewer still did both.

The failure of these studies to produce significant results is thought to signify one of several possibilities: (1) either there is no production function in schools, or the production function for each child is so different that no overall conclusion may be drawn; (2) the researchers engaged in these studies have either failed to control for variables that "wash out" the analyses, or have studied the wrong variables; (3) school inputs do not affect school outputs, and so achieving fiscal equity is not necessary; (4) school inputs do matter, but

not in determining achievement test scores--rather in more important attitudes and values and in eventual job income and status.

Each argument has its merits. Murnane (1980) suggests that the results of a decade of research on school productivity must be construed as demonstrating that no production function exists, at least not in an identifiable form that is uniformly applicable to all settings. Similar findings exist on a micro-level in the classroom, where studies have shown individual learner preferences with resultant differences in achievement test scores (Brophy, 1970).

Researchers have been quite remiss in controlling for the effects of socio-economic status and previous student achievement. Both are generally considered to be non-school variables with considerable ability to mask the effects of school-related variables. Walberg and Bargen (1977) found only 3 of 21 studies adjusted for the effects of both student native ability and parental background. Contradictory findings in light of such imprecise studies are unsurprising.

Even if intervening variables were controlled, a great deal of concern must be directed toward the independent variables chosen for inclusion in the analyses. Researchers are forced to deal with surrogate measures of quality and quantity. Thus, teachers who are older, with more experience and higher degree status, are assumed to reflect higher quality. Per pupil expenditures, teacher characteristics, even the dependent variable, achievement tests, are all expressed by means or medians which may well disguise relevant interactions. Appropriate variables, at the level of the child, are simply not readily available, leading to ecological fallacies (Brischetto, 1980).

Much has recently been made of the failure of the school finance reform movement to achieve reduction of intra-state per pupil expenditure disparities (Brown et al., 1978). Other reformers have concentrated on the distribution of funds to minorities (Brischetto, 1980). However, equitable fiscal policy is meaningless if alternative learning outcomes are not achieved. Unless the presence or absence of differential funds can be demonstrated to affect achievement test scores

uniformly in all settings, policy makers cannot be expected to direct funds toward disparity reduction. Indeed, Goertz (1979) reports the increased revenues received by poor wealth, low expenditure school districts being used to lower tax rates, rather than increase expenditures. Such concerns may also be the genesis of what econometricians term "price" or the willingness of school districts to respond to state aid programs (Feldstein, 1975; Ladd, 1975). Other investigators suggest that insufficient attention has been directed toward intra-school distribution, rather than inter-school allocations (Monk, 1980). Unless it can be demonstrated that money affects test scores, intra-state disparities, disparities between minorities, and intra-school disparities are of dubious concern.

Finally, other investigators (Walberg, 1979) suggest that money might matter, but not directly in achievement test scores. Instead, there may be an intervening variable, such as school environment, that is affected by school inputs, which in turn affects achievement. Thus, increases in school expenditures will have positive effects upon the school learning environment and student attitudes, which may or may not raise achievement.

Earlier work in New York state had been conducted by Goodman (1959) with a reanalysis by Kiesling (1967). Iowa test scores of 70,000 7th and 11th graders in 102 school districts were analyzed, and student achievement was positively related to expenditures per pupil, particularly in districts above 2,000 pupils. For small rural districts, however, the relationship was random and sometimes negative. Neither Goodman nor Kiesling held constant the effects of both socioeconomic status and previous test achievement, and their sample was limited to one-seventh of the school districts in the state.

The extant work was undertaken to confirm Goodman's and Kiesling's findings in New York state, with confounding variables held constant and the entire population of public school districts and children at certain grade levels. The remainder of the article will describe previous literature, explain the methodology used, and discuss the findings.

2. Previous Findings

More than a decade before the Coleman Report was to shake the foundations of the educational establishment, researchers were interested in the relationship between expenditures per pupil and the quality of education (Gauerke and Childress, 1967). Typically, their studies consisted of the relationship between expenditures and test scores, or process variables, such as innovativeness. However, these studies were seriously flawed, in that they controlled neither for student background or native ability. Thus, it could be argued that student SES and previous achievement, absent high expenditures, would produce the same results. (Gutherie, 1970)

By 1956, educators were becoming more sensitive to this criticism and began to include measures of socioeconomic status. One study involved a nationwide sample of 9,000 ninth grade students in 100 schools and 8,357 twelfth graders in 106 schools. Principals' responses were obtained as a measure of SES. Although uncertain of the effectiveness of their surrogate measure in preventing contamination of their results, they reported some school variables to be related to achievement, including expenditures per pupil (Mollenkopf and Melville, 1956).

In 1959, Goodman used a sample of 102 school districts, with 70,000 7th and 11th graders in New York state. Controlling for the SES of parents, he reported that for 7th graders, per pupil expenditures and teacher experience were significantly correlated with achievement. A 1962 nationwide study of 206 10th and 12th graders in 46 states (Project Talent data) with 1960 Census data resulted in a significant correlation between beginning teachers' salaries and teachers' experience and students' test scores (Thomas, 1962). Kiesling undertook a reanalysis of Goodman's data in New York state in 1965, and found that for large school districts (2,000) expenditures per pupil and student achievement are significantly related. For small, rural districts, the relationship was random, even negative (Kiesling, 1965). Benson (1965) in a study of 249 California school districts, with SES controlled, found teachers' salaries and expenditures per

pupil to be positively related to pupils' achievement.

But these positive findings between expenditures per pupil, other school characteristics, and student achievement were soon cast in doubt, primarily by the work of non-educators. The sociologists who investigated this relationship believed that students' achievement was highly correlated with socioeconomic status, and undertook analyses to adequately remove this confounding influence from school effects upon achievement. Much to the astonishment of educators, they found ". . . that schools bring little influence to bear upon a child's achievement that is independent of his background and general social context." (Coleman, 1966)

The furor that followed this finding, with charges of inadequate measures, inappropriate statistical treatments, and improper manipulation of the measures, spawned extensive reanalyses. Armor (1972) explicitly shaped his methodology to not exaggerate family effects to the detriment of school effects, but came to the same conclusions as Coleman. Jencks (1972) also concurred.

More negative evidence came from Averch (1974) who surveyed all of the studies in five areas: input-output; process; organizational; evaluation; experiential. He reached three policy implications:

> "1. Research has not identified a variant of the existing system that is consistently related to the students' educational outcomes.
> 2. Increasing expenditures on traditional educational practices is not likely to improve educational outcomes substantially.
> 3. There seem to be opportunities for significant reduction or redirection of educational expenditures without deterioration in educational outcomes." (p. 173)

Mayeske's (1969) study also consisted of a reanalysis of Coleman (et al., 1966). He found only 36 percent of the differences in student achievement attributable to schools, the rest to within-school and non-school factors. He concludes that school effects will be seriously underestimated if achievement data are first corrected for student socioeconomic status. Thus, the present state of our quantitative models will not permit untangling of the effects of home,

school, and environment upon student outcomes.

Other studies of school effectiveness (Gutherie, 1970; Suppes and Wells, 1974) have demonstrated contradictory findings. In their analysis of 17 studies, Suppes and Wells found teachers' experience significant in 8 studies and insignificant in 10, per pupil expenditure significant in 5, and insignificant in 4, teacher education significant in 2, and insignificant in 5. Gutherie (1970) found similar results in his review of 19 studies: per pupil expenditures were significant in 7; teachers' salary in 6; teacher experience was significant in 7; teachers' degree level in 2; school size in 3.

In a more optimistic review of 23 studies, 21 of which controlled for socioeconomic status and 4 previous student achievement (and 3 both), Bargen and Walberg (1977) found teacher salary significant in 7, with expenditures per pupil significant in 3 others. See Table 13-1. Teacher experience and achievement was significant in 11, and teacher education and achievement was significant in 7 studies. In their own study of the Chicago School system (1978), Bargen and Walberg found teacher quality and achievement related, but in a curvilinear manner.

Thus, it appears that prior to the application of socioeconomic status to control for home background, investigators found significant relationships between expenditures per child, school characteristics and achievement. These findings either washed out, or became unstable when researchers began to uniformly apply SES measures as controls. Studies of New York state seem to run contrary to this generalization, however, and this study addresses this issue.

Table 13-1. Major studies of teacher quality and class size.

This table was removed prior to submission to the ERIC Document Reproduction Service because it is copyrighted.

Refer to: Bargen, M. and Walberg, H. J. "School Performance". In: Herbert J. Walberg (ed,), Evaluating Educational Performance, Berkeley, McCutchan, 1974, pp. 239-254.

3. Methodology

All data was supplied by the New York State Department of Education on a data tape. The file was composed of the 1975-76 Consolidated Data Base and the 1975-76 Finance Tapes for all 705 school districts in the state. A complete list of variables is included in Appendix A and was associated with an earlier study (Fowler and Furno, 1978). Variables of interest in this study were: State pupil evaluation tests of reading and mathematics conducted by the State for all districts in grades 3, 6, and 9; Regent's tests and school characteristics, such as per pupil expenditure, average teacher degree level, etc. Table I lists the dependent and independent variables used in the study.

The data file supplied by the New York State Education Department was then merged with 1970 Census data for New York school districts obtained from the U.S. Bureau of Census 5th Count data tape. Using a methodology devised by Charles Nam in 1963, the median education, occupation, and income of each district was transformed into a continuous scale from 0 to 100 for each measure (Working Paper No. 15, 1963). Matches were obtained for 642 districts (of 705). Since 1963, Nam (et al., 1975) have been engaged in an ongoing study of the dynamics of occupational status stability and change. Occupational status scores, income, and education scores that were appropriate for the 1970 Census were used. Nam (1975) comments upon the methodology for occupations, which was similarly applied to education and income.

> For each of the categories, the procedures for deriving the scores are essentially the same, being based on the combined average levels of education and income for incumbents of the occupation. More specifically, the procedure involved (a) arraying occupations according to the median educational level of persons in the experienced civilian labor force, (b) arraying the occupations separately according to the median income level of each occupation, (c) by using the number of persons engaged in each occupation as weights, determining the cumulative interval of persons in each occupation for each of the two arrays, and (d) averaging the midpoints of the two cumulative distributions of occupants and dividing by the experienced civilian labor force to get a status score for the occupation. The

> resulting scores can take values between 0 and 100, and a score indicates the approximate percentage of persons in the experienced civilian labor force who are in occupations having combined average levels of education and income below that for the given occupation.

Nam completed a major re-classification of the occupational list in 1970 to increase the homogeneity of the detailed occupations and to reduce the number of persons listed under occupations not classified elsewhere. Examination of the scores between censuses demonstrates little change for the vast majority of occupations, although notable changes occurred for a small group of occupations.

The computer program, Statistical Package for the Social Sciences (Nie, et al., 1975) was utilized to obtain means, frequencies, and skewness of the variables of interest. Variables displaying positive skewness in excess of 1.00 were transformed into natural logarithms, reducing the abnormality of the distribution. Per pupil expenditure levels (ranging before transformation from $800 to $4,000) were treated in this manner, as was property wealth, enrollment, total professional staff/ADA total classroom teachers/ADA other professional staff, teachers with doctorates, and instructional salaries/ADA.

Relationships between nine dependent and 15 independent variables were examined. The dependent variables were state PEP tests of reading and mathematics for all districts (705) for grades 3, 6, and 9, the percent of students scoring two years below certain achievement levels for those grade levels, and the average Regent's score for each district. The fifteen independent variables consisted of school district characteristics (see Table I).

The following procedure was used: the average SES score and previous grade test score were regressed upon the dependent test score of interest such as the median 9th grade reading score in a district. These regressions provided metric weights that were used in an equation to estimate the residual of dependent test score controlling for previous grade level test score and SES.

TABLE I

Variables of Interest: Aggregated at School District Level

Dependent Variables

1. School District Median PEP 9th Grade Reading Score
2. " " " " 9th Grade Math Score
3. " " " " 9th Grade Below Minimum Competency Reading Score
4. " " " " 9th Grade Below Minimum Competency Math Score
5. " " " " 6th Grade Reading Score
6. " " " " 6th Grade Math Score
7. " " " " 6th Grade Below Minimum Competency Reading Score
8. " " " " 6th Grade Below Minimum Competency Math Score
9. School District Average Regent's Score

Independent Variables

1. Expenditure Per Pupil (transformed into natural logarithm)
2. Enrollment (" " " ")
3. Total Professional Staff/WADA (" " " ")
4. Total Classroom Teachers/WADA (" " " ")
5. Other Professional Staff/WADA (" " " ")
6. Tax Rate for Education
7. Number of Teachers with BA
8. " " " " BA + 30
9. " " " " MA
10. " " " " MA + 30
11. " " " " Doctorate (transformed into natural logarithm)
12. Mean-Years of Teaching Experience in District
13. Mean-Age of Teachers in District
14. Mean-Instructional Salaries/WADA (transformed into natural logs)
15. Total State Aid Per Total Aidable Pupil Unit (TAPU)

The computation of the metric weights was obtained by regressing the variable to be controlled against the variable of interest. The predicted value of the variable of interest was then computed using formula 1.1:

(1.1) Est 1 = constant + (b controlled variable)

The residual of the variable of interest was obtained by subtracting this estimate from the actual value of the variable, using formula 1.2:

(1.2) Res 1 = (Variable of interest - Est 1)

Scattergrams were then plotted against Res 1, which display the correlations with the effects of the controlled variable removed. (Walbert & Rasher, 1976)*

Table II contains the results of the regressions of SES and previous grade level test score upon the upper grade level test score. As might be anticipated, the previous grade level score predicts the upper grade level score quite well, entering the step-wise regression first, and in the case of 9th grade reading, 6th grade scores explained 38 percent of the variance. Including SES explained another 3 percent of the variance. The F ratios are significant at the .01 level. Similar results are obtained for 9th grade math scores. The high zero-order correlations (.43, .44) between the test scores and SES somewhat mask the true effects of SES. Dropping the lower-grade level test from the equation significantly enhances the explanatory power of SES. For the Regent's score, SES explained 26 percent of variance, ($p < .01$).

It was hoped that this procedure would permit the assessment of the effects of school characteristics upon median district test scores without the confounding effects of SES and differential native ability cited by previous researchers in school effectiveness. Both measures are rather undesirable surrogates. Ideally, the socioeconomic measure should consist only of the grade level taking the PEP tests, not the entire district's SES. Similarly, it would be preferable to have either the longitudinal scores of the students of a grade level from three years hence, or an I.Q. measure. Unfortunately, only the same year's

*I am indebted to Herbert J. Walberg, Research Professor of Human Education, University of Illinois at Chicago Circle, for his suggestions regarding this technique.

TABLE II

Regression of Previous Grade Level Test Score and SES Upon Upper Grade Level Score

Dependent Variable: 9th Grade PEP Reading

Independent Variables	Multiple R	R^2	Simple R	b	Beta	DF	F
PEP 6th Reading Score	.61845	.38248	.61845	.4192	.5317	2	233.58
SES	.64211	.41231	.43190	.8003	.1932	608	30.85

PEP test results were available for an earlier grade level. Despite these considerable flaws, it is remarkable that the surrogate measures were as highly related to the criterion variable as the results indicate.

Pupil Evaluation Program (PEP) tests are required of all pupils in grades 3 and 6 in the public and non-public schools in New York state. They are also offered on an optional basis for grade 9 pupils. The tests are reading and mathematics tests designed to measure instructional goals promulgated by NYSED. Since 1973, mathematics tests have been available in a Spanish-language form. Non-English speaking pupils are not given the reading tests. The same tests are administered without change over a 5-year period. New forms for the reading test were used in 1973, and in math, in 1969.

In addition, students whose scores fall below a certain point (the lowest 3 stanines) of all pupils tested in each grade in 1966, are identified for purposes of identifying students toward whom ESEA Title I funds should be directed. These students are captured by the variable "Percent Below Minimum Competency."

Reliability coefficients using Kuder-Richardson Formula 21, as well as split-half reliability coefficients, have been obtained for grades 3 and 6 for both reading and math. Table III records the coefficients and standard errors obtained by each method.

Regents examinations are end-of-course high school achievement tests, offered three times a year in 25 different high school subjects. Only students enrolled in "Regents" type courses take the tests, and these are typically college-bound students. The Regents Scholarship and College Qualification Test (RSCQT) consists of two parts, administered in the morning and afternoon. Part I is a test of general scholastic aptitude. Part II is a test of subject matter achievement. No data in the reliability of the RSCQT was available at the time this article was written.

TABLE III

Kuder-Richardson Reliability Coefficients and Standard Errors of Measurement for the Total and Part Scores of the Mathematics Tests

PART	Grade 3				Grade 6			
	M	S.D.	r	SE_m	M	S.D.	r	SE_m
Concepts	13.4	5.4	.81	2.4	12.1	5.4	.80	2.4
Computation	7.2	4.1	.81	1.8	9.1	4.0	.73	2.1
Problem Solving	10.3	5.0	.87	1.8	9.1	4.4	.78	2.1
Total	31.1	13.4	.93	3.5	30.4	2.6	.91	3.8

Split-Half Reliability Coefficients and Standard Errors of Measurement for the Total Scores of the Mathematics Tests

Test	N	M	S.D.	r	SE_m
Grade 3	7,576	30.1	12.8	.93	3.4
Grade 6	7,934	30.9	12.3	.92	3.5

Intercorrelations of Part Scores for the Mathematics Tests

Grade	Part	Computation	Problem Solving	Total Test
3	Concepts	.73	.75	.92
	Computation		.69	.87
	Problem-solving			.90
6	Concepts	.63	.69	.89
	Computation		.64	.84
	Problem-solving			.89

Kuder-Richardson Formula 21 Reliability Coefficients and Standard Errors of Measurement for the Tests in Reading

Test	N	Mean	Sigma	r	SE_m
Grade 3	299,132	26.9	12.5	.94	3.08
Grade 6	308,320	31.4	11.1	.92	3.06

Split-Half Reliability Coefficients and Standard Errors of Measurement for the Tests in Reading

Test	N	Mean	Sigma	r	SE_m
Grade 3	7,276	27.2	11.9	.94	2.94
Grade 6	7,934	32.1	10.5	.93	2.80

*Bureau of Elementary and Secondary Educational Testing, State Education Department, Albany, New York 12224

* Reprinted with permission of Winsor A. Lott, Chief, Bureau of Elementary and Secondary Educational Testing.

4. Results

Table IV displays in summary form the results of the scattergrams between the dependent score and the fifteen independent variables, with SES and previous grade level test score controlled. The relationship between expenditure per child and 9th and 6th Reading and Math PEP median scores and average Regents score was negative, and significant at the .01 level for three of the five variables. The correlations were extremely low (ranging from -.04 to .21). Indeed, when results for the 4 dependent variables "Percent Below Minimum Competency" are excluded, all of the significant independent variables were negatively correlated with the dependent variables, at similarly low levels, with the single exception of number of teachers with doctorates. Contrary to Kiesling's (1967) and Goodman's (1959) studies, neither expenditure per child nor school district size had positive correlations with test scores.

Observing the column "# Positive Corr." (Table IV) is instructive. Tallied in that column are the number of positive correlations (not necessarily significant) obtained between each of the fifteen independent variables with that row's dependent variable. Thus, for the 9th grade PEP Reading Score, only 4 independent variables correlated positively (1 significantly, $p < .05$) with the dependent variable. The median correlation (-.20) reflects this finding.

Notice, however, that when one considers only "Percent Below Minimum Competency," that is, the percentage of students scoring below the minimum competency score (two years below grade level), the signs of the coefficients quickly turn positive. For three of the four test scores, 11 positive correlations were found among 15 independent variables, and 9 of 15 for the fourth. Further, both expenditure per pupil and enrollment were positively correlated (6th grade math, +.22, and $p < .01$, was one of the highest correlations obtained).

The findings for the average district Regents score appear about the same as for 3rd and 6th grade median scores, that is, only 5 positive correlations were

TABLE IV. RESULTS OF SCATTERGRAMS WITH SES AND PREVIOUS GRADE LEVEL TEST SCORE REMOVED

Dep. Vars. / Ind. Vars.	Log of Expend. level	Log of enrollment	Tax Rate for education	Log - Total Prof. Staff/WADA	Log - Total Class Room Teachers/WADA	Log - Other Prof. Staff/WADA	# of Teachers with Bachelor's	# of Teachers with Bachelor's + 30	# of Teachers with Master's	# of Teachers with Master's + 30	Log of # of Teachers with Doctorates	Years Teaching in District 50th Quartile	Teacher's Age 50th Quartile	Log - Instruct Salaries/WADA	Total State Aid per TAPU	# Positive Corr.	Median Correlation
PEP 9th Grd Rng	-.07*	-.08*	-.08*	-.06*	-.07*	-.007	+.05	+.05	-.05	-.10**	+.13*	-.05	-.01	-.10**	+.02	4	-.20
PEP 9th Grd Math	-.11**	-.12**	-.13**	-.06	-.03	-.11**	+.03	+.10**	-.01	-.12**	+.17**	-.04	-.02	-.13**	+.06	4	-.050
PEP 9th% BEL Rng	+.13**	+.14**	+.13**	+.07*	+.05	+.10**	-.06	-.07*	+.03	+.14**	-.09	+.06	+.03	+.14**	-.09**	11	+.055
PEP 9th% BEL Math	+.07	+.13	+.09**	+.04	+.03	+.04	-.07*	-.05	+.04	+.11**	-.09	+.02	+.03	+.10**	-.02	11	+.035
PEP 6th Grd Rng	-.21**	-.20**	-.20**	-.10**	-.08**	-.12**	-.17**	+.06	-.09**	-.18**	+.14**	-.03	-.06	-.23**	+.11**	3	-.11
PEP 6th Grd Math	-.06	-.13**	-.12**	-.06	-.06*	-.008	+.05	+.05	-.03	-.04	+.14**	+.09**	+.05	-.06	+.04	6	-.035
PEP 6th% BEL Rng	+.05	+.14**	+.12**	+.06	+.04	+.05	-.02	-.05	+.003	+.03	-.15*	-.09**	-.05	+.06*	-.01	9	+.035
PEP 6th% BEL Math	+.22**	+.20**	+.17**	+.15**	+.12**	+.14**	-.15**	-.04	+.08*	+.17**	-.13*	+.05	+.08*	+.23**	-.12**	11	+.10
AVER Dist Regent's Score	-.04	+.07*	-.06	-.07*	-.07*	-.04	-.04	+.06	+.02	-.02	+.09	+.03	-.003	-.03	-.003	5	-.025
#Positive Corr.	4	5	4	4	4	4	3	5	5	4	5	5	4	4	4		
Median Corr.	-.04	+.07	-.06	-.06	-.03	-.007	-.04	+.05	+.003	-.02	+.09	+.02	-.003	-.03	-.003		

**Significant at .01 level
*Significant at .05 level

obtained, with only one significant finding for enrollment. The low correlation, however, (+.07, $p < .05$) indicates how weak this relationship is. The median correlation for all fifteen independent variables is a -.025, indicating the absence of a relationship.

A final consideration are the rows below the nine dependent variables, where again the number of positive correlations and the median correlation for each of the independent variables is tallied. Note that the number of positive cases never exceeds chance, and that virtually all of the positive relationships refer to the "Percent Below Minimum Competency" dependent variables. Thus, expenditure level per child demonstrates a negative relationship with 3rd and 6th grade median PEP reading and math scores, and with the average Regents score.

5. Discussion

The results generate two conclusions, both of which deserve extended discussion. First, the overall findings again confirm the puzzling, inconsistent pattern between achievement test scores and school characteristics, and even demonstrates negative correlations between supposed quality variables (such as expenditures per pupil) and achievement test scores. Second, the results also suggest an interesting hypothesis: that additional expenditures per child and other quality-related measures raise the test scores of the poorest performing students.

There has always been extended debate about the research on school productivity. The conclusion drawn by researchers summarizing studies of school productivity has been that the differences in students' achievement have not been attributable to differences in educational environments. Such findings have not ended the search for the unusually effective school and its driving mechanism.

> ". . .rather than closing up shop or changing their business, educational researchers and school planners have raised issues of research methodology and interpretation to argue that conclusions on the ineffectiveness of schools are premature or misleading." (McPartland and Kurweit, 1979)

A number of arguments are advanced to explain findings similar to those presented here. First, studies of resource allocation and school productivity have typically sampled natural occuring variations of schooling and student populations, and their concomitant non-school influences (Coleman, 1966; Mosteller and Moynihan, 1972; Jencks et al., 1972; Hauser, 1971; Averech et al., 1972; Spady, 1973). The results of these studies indicate that school differences account for little of the inter-school variations in achievement test scores. Because these studies were conducted in a non-controlled setting, typically ex post facto, two objections are raised to the methodology: the natural variations are insufficient to reveal meaningful relationships; the separation of school and non-school

influences on student achievement has not been adequately attained with current methodology.

The methodology used in the extant study was designed explicitly with this objection in mind. Non-school influences were removed from the analysis by controlling for the effects of socioeconomic status and previous achievement. However poor these surrogates are for family teaching and student native ability, they did, in combination, explain about 40 percent of the variance in median district test score. Thus, the school effects shown in Table IV may be interpreted as having the confounding effects of non-school effects removed. As other researchers have found, school effects are small and inconsistent. Armor (1972), of course, conducted his reanalysis of the Coleman study precisely for the opposite rationale, to detect whether or not Coleman, in excluding SES, had eliminated that portion of the variance associated with both school factors and student background, thus exaggerating the influence of home factors and understating the influence of school characteristics. He concluded that alternative methodologies yielded approximately the same results.

An alternative interpretation of studies asserting that school differences do not account for inter-school achievement variation is that school environments, far from being weak and ineffective, are so uniform and effective and schooling so similar, that the greater variability of non-school variables will correlate more highly with differences in student outcomes (McPartland and Kurweit, 1979). A similar argument is that the naturally occurring range of school contrasts is so small that the relationships are correct, but that the relationships would be dramatically different if the range of contrasts were only greater. Class size is often mentioned in this regard.

A look at the data, however, obviates these explanations. District per pupil expenditure, before natural log transformation, ranged from 834.88 to 4393.64, with a standard deviation of 466.28, and was skewed positively, 2.03. By contrast, district SES had a range of only 32.8, and a standard deviation of 5.9.

Other investigators have produced more telling arguments for the results obtained here. Clearly, the measures of school resources are distal proxies in this study, and there is little reason to believe that the school district attributes routinely collected by the State Department of Education reflect quality measures that would affect students' test scores (Erikson, 1977). Another concern is that the measures used here are aggregated at the school district level which may mask within-school educational differences that have impact on student achievement (Bowles and Levin, 1968). Hanushek found differences between school averages of faculty characteristics and individually measured faculty characteristics and student outcomes (1970). Some investigators have insisted that the duration of the students' exposure to resources and the differential distribution of school level inputs within schools must be studied to determine schooling outcomes. Early studies of this phenomena have proved less than promising, however (Karweit, 1976). Further, although intra-district distribution is differential, it appears to be based upon tested verbal ability rather than SES or some invidious discrimination (Heyns, 1974; Summers and Wolfe, 1976).

A final argument to explain the failure of input-output studies to discover linkages between resources and achievement has been to critique the outcome measures as inappropriate, insensitive, or irrelevant. In answer to criticisms that standardized tests are inappropriate to measure the cognitive outcomes of schools, Cooley and Lohnes (1976) argue that although such tests cannot measure the behavioral objectives of specific programs, they provide the best measure of general educational development. Madaus et al., (1979) assert that ". . . it seems odd to measure what is admittedly a side effect of education while at the same time ignoring the more direct results of particular curricula and courses." (p. 209) In order to assess the relative advantages of one measure over the other, Madaus et al., compared standardized and curriculum-based measures of cognitive attainment at the individual classroom unit of analysis, finding the between-class variance explained by curricular tests larger than that explained by standardized

tests. Others assert that there may be other student outcomes of schooling that have not even been considered in the research on school effects for which measures are currently unavailable (Coleman, 1972; McPartland and Karweit, 1979).

Three student measures were utilized in this study. PEP reading and mathematics median school district scores and school district average Regents score. Whatever arguments one wishes to make concerning the relationship of generalized achievement tests to curricular tests, a reasonable expectation is that reading and mathematics skills are desirable outcomes for high school students and are reasonably proximate to curricular tests. Further, the Regents exam is composed of two parts, the first being a general scholastic aptitude test and the second being a subject matter achievement test. It would appear that the average score would then reflect attainment in both areas, silencing much criticism of previously used measures (albeit a less aggregated measure would have been a preferable unit of analysis). Further, since it is usual to assert that removing SES from the equation results in removing the school effects that are covariant with student background factors (i.e., that the extraordinary skills of teachers' effects on high SES students' achievement are removed), it is desirable to have a measure of Regents examinations. The desirability of this measure lies in the fact that Regents' examinees typically are the top ability groups in a school and may be thought of as a correction for school effects when removing SES.

The second major finding of this study is that additional expenditures per child and other quality related measures are effective in raising the test scores of the poorest performing students. For the dependent variables "Percent Below Minimum Competency" (i.e., students two years behind certain achievement levels for a given grade level) of the thirty independent variables that proved significant for the four sets of dependent variables, twenty-four were positive, but the correlations were still weak (from +.06 to +.22). For these students, the effects are something akin to the traditional commitment that expenditures affect school attributes, which in turn affect student achievement. For example, the correlations

between poorly achieving students and numbers of teachers with B.A.'s turns negative (significantly two of four times), but positive for the variable "Teachers with M.A." Expenditure per child, district size, and tax rate for education are all positively correlated with the scores of poorly achieving students. Such findings are consistent with those of Brown and Saks (1975) who demonstrated that additional expenditures reduced the standard deviation of test scores within each district. Thus, it would appear that those who argue that money makes a difference are somewhat justified: the poorest achieving students are aided.

What policy conclusions flow from these results? First, much has been made of the failure of production function studies to find uniform effects. Given the distal surrogate measures that typically have been utilized, without an ability to adequately separate within school and between school variance, the findings are hardly surprising. A redirection of research on educational effects by both selecting samples to maximize the range of school environment variation and to utilize more sensitive measures would seem a reasonable first step. Second, much has been made of the continuing and widening expenditure disparity between property-wealthy and property-poor school districts (Brown et. al., 1977). Attempts by states to alter this relationship have not been successful as property-poor school districts utilize additional revenues to lower their tax rates, rather than raise expenditures per pupil. (Goertz, 1978) Libertarians applaud this as consistent with school district choice. Equalitarians would decry such place discrimination because they believe the results of unequal expenditures inevitably means differences in the quality of education.

Oddly, the results lend support to both sides. Expenditures per pupil, and the characteristics typically purchased with them, do not appear to affect median district test scores. Thus, disparities in expenditures are not of concern if one only attends to median scores. But for equalitarians, the results demonstrate the necessity of reducing the disparity in per pupil expenditures if one wishes to

raise the test scores of students currently performing at least two grade levels below their peers. If, indeed, replication confirms the effect of increased expenditures per child upon the lowest stanines, current proponents of district power equalizing and other wealth neutral school finance reforms that do not create equality of expenditures may have to reformulate their positions.

Third, the current debate on minimum competency testing for the award of a high school diploma should intensify research on the effects of additional funds or school resources on the poorest scoring students. There is reason to believe that the current intra-school resource distribution system is inequitable in a manner designed to perpetuate class or ability stratification (Blau and Duncan, 1967; Hanser, 1969; Heyns, 1974). If money and staff attributes matter, however, the elixir for failing students is deceptively simple.

Finally, this study, unlike Brown and Saks (1975) did not consider the range of standard deviations associated with school district attributes. While it is encouraging to suggest that poorly achieving students may be aided by certain kinds of school profiles, entirely different concerns arise if, as Brown and Saks suggest, the entire range of scores is reduced when the lowest performing students are attended to. If, indeed, reducing the expenditure disparity between school districts not only has positive implications for the lowest achieving students, but negative effects for the highest achieving students, those of us concerned with obtaining equality of educational opportunity may find our task Sisyphean.

REFERENCES

Armor, D. J. "School and Family Effects on Black and White Achievement: A Re-examination of the USOE Data." In: F. Mosteller and D. P. Moynihan (eds.) On Equality of Educational Opportunity. New York: Random House, 1972, pp. 168-229.

Averch, H. et al. How Effective is Schooling? A Critical Review and Synthesis of Research Findings. Santa Monica, California: Rand Corp., 1974.

Benson, C.S. et al. State and Local Fiscal Relationships in Public Education in California. Senate of the State of California, (March, 1965).

Blau, P. M. and Duncan, O. D. The American Occupational Structure. New York: Wiley, 1967.

Bowles, S., and Levin, H. "The Determinants of Scholastic Achievement - An Appraisal of Some Recent Evidence." Journal of Human Resources 3 (1968): 3-24.

Brischetto, R. Minorities, the Poor and School Finance Reform of the 1970's: A Summary of Findings in Six States. Paper presented at the American Education Finance Association National Conference, March 16-18, 1980, in San Diego, California.

Brophy, J. and Good, T. "Teachers' Communications of Differential Expectations for Children's Classroom Performance: Some Behavioral Data." Journal of Educational Psychology 61 (1970), 356-374.

Brown, B. W. and Saks, D. H. "The Production and Distribution of Cognitive Skills Within Schools." Journal of Political Economy, 1975, Vol. 83, No. 3, pp. 571-593.

Coleman, J. S. et al. Equality of Educational Opportunity. Washington, D.C.: U. S. Government Printing Office, 1966.

Cooley, W. W. and Lohnes, P. R. Evaluation Research in Education. New York: Wiley, 1976.

Feldstein, M. S. "Wealth Neutrality and Local Choice in Public Education," American Economic Review 65 (March, 1975): pp. 75-80.

Gauerke, W. E. and Childress, J. R. (eds.) The Theory and Practice of School Finance. Chicago: Rand McNally Co., 1967, pp. 279-308.

Goertz, M. E. — Where Did the 400 Million Dollars Go? The Impact of the New Jersey Public School Education Act of 1975, Princeton, New Jersey: Educational Testing Service, 1978.

Goodman, S. M. — The Assessment of School Quality. Albany: The State Education Department of New York, 1959.

Gutherie, J. W. — "A Survey of School Effectiveness Studies." In: Do Teachers Make a Difference? Washington, D.C.: U.S.O.E., 1970, pp. 25-54.

---------- et al. — Schools and Inequality. Cambridge, Mass. MIT Press, 1971.

Hauser, R. M. — Sociometric Background and Educational Performance. Washington, DC: American Sociological Association, 1971.

Heyns, B. — "Social Selection and Stratification Within Schools." American Journal of Sociology 79 (1974): 1431-34.

Jamison, D., Suppes, P. and Wells, S. — "The Effectiveness of Alternative Instructional Media: A Survey." Review of Educational Research 44 (1974): 1-67.

Jencks, C. S. and Brown, M. D. — "Effects of High Schools on their Students." Harvard Educational Review 45 (1975): 273-324.

Jencks, C. et al. — Inequality: A Reassessment of the Effects of Family and Schooling in America. New York: Basic Books, 1972.

Karweit, N. — "A Reanalysis of the Effects of Schooling on Achievement." Sociology of Education 49 (1976): 237-246.

Kiesling, H. J. — "Measuring a Local Government Service: A Study of School Districts in New York State." Review of Economic and Statistics, Vol. XLIX, No. 3. (August, 1967), pp. 356-367.

Ladd, H. F. — "Local Education Expenditures, Fiscal Capacity and the Composition of the Property Tax Base." National Tax Journal 28 (June, 1975), pp. 145-148.

Madaus, G.F. et al. — "The Sensitivity of Measures of School Effectiveness." Harvard Educational Review (May, 1979) 49; 2: 207-230.

Mayeske, G. et al. — A Study of Our Nation's Schools. Washington, DC: U. S. Office of Ed., 1969.

Mollenkopf, W. G. and Melville, S. D. "A Study of Secondary School Characterisitcs as Related to Test Scores." Research Bulletin 56-6. Princeton: Educational Testing Service, 1956.

Mosteller, F. and Moynihan, D. P. (eds.) On Equality of Educational Opportunity. New York: Vintage Books, 1972.

Murnane, R. J. "Interpreting the Evidence on School Effectiveness." Paper presented at the American Education Finance Association National Conference, March 16-18, 1980, in San Diego, California.

McPartland, J. M. and Karweit, N. "Research on Educational Effects." In: Walberg (ed.) Educational Environments and Effects. Berkeley, California: McCutchan, 1979, pp. 371-385.

Nam, C. B. et al. Occupational Status Scores: Stability and Change. Proceedings, 1975 American Statistical Association.

Spady, W. G. "The Impact of School Resources on Students." In: Review of Research in Education, Vol. I, Frederick N. Kerlinger (ed.), Itasca, Ill.: F. E. Peacock, 1973, pp. 135-177.

Summers, A. A. and Wolfe, B. L. "Which School Resources Help Learning?" IRCD Bulletin 11, No. 3, 1976.

Thomas, J. A. "Efficiency in Education: A Study of the Relationship Between Selected Inputs and Mean Test Scores in a Sample of Senior High Schools." Unpublished Ph.D. dissertation (Stanford University: School of Education), 1962.

U. S. Bureau of the Census, Methodology and Scores of Socioeconomic Status, Working Paper No. 15, Washington, DC., U. S. Government Printing Office, 1963.

Walberg, H. J. and Rasher, S. P. "Improving Regression Models." Journal of Educational Statistics. Vol. 1, No. 3. (Autumn, 1976), pp. 253-277.